HW#1 --- last modified Wednesday, 31-Jan-2018 12:44:37 PST.

[Solution set](.).

Due date: Feb 12

Files to be submitted:
  Hw1.zip

Purpose:To refresh our memories about CS157a and to experiment with how DBMS implementation can influence the design and deployment of databases.

Related Course Outcomes:

The main course outcomes covered by this assignment are:

CLO 1 -- To understand how a DBMS systems is physically implemented.

Specification:

This homework will consist of two parts, a written part and a coding part. Both parts will be submitted in the Hw1.zip file. The written part should be in a file Hw1.pdf. For this part of the homework you should write solutions for the following questions. In what you turn in, first copy and paste the question, then write your solution to it beneath it.

1. Suppose SuperHW drive has the following characteristics:
   - There are 3 platters with 150,000 tracks each.
   - Tracks have 4000 sectors of 1024 bytes each.
   - 10% of each track is used for gaps.
   - The disk rotates at 5400 RPM.
   - The time it takes the tape head to move n tracks is $1 + 0.0002n mS$.

   Answer the following questions about the SuperHW drive:
   1. What is the capacity of the disk?
   2. Suppose this disk a 2.5in laptop drive, what is the average bit density in the sectors of the outermost track?
   3. What is the maximum seek times?
   4. What is the average rotational latency?
2. Suppose we are scheduling I/O requests for the SuperHW drive given in problem 1, and the head is initially at track 75,000. Further, suppose requests come in according to the table below:

Cylinder Requested	Time Request Made From Start Time in mS
1000	0
50000	2
8000	11
44000	19

At what time is each request serviced fully if: (a) the elevator algorithm is used (initially moving inward)?

(b) We use first-come-first-served scheduling.

3. Suppose we have 3 data drives and 1 redundant disk and we are using RAID level 4. Assume blocks are a single byte. (a) What would be the redundant block corresponding to the three data blocks: 10111011, 11000011, 01010101? (b) What would be the redundant block corresponding to the three data blocks: 10000011, 10010011, 11011101?

The coding part of the assignment assumes you have a version of Mysql/MariaDB installed on your computer and have Sqlite 3 installed. If you don't already have these DBMSs installed probably the easiest way to get them is probably to use a package manager. On Linux, you can use the default manager **apt**, on a Mac, I recommend homebrew, and on a PC, I recommend chocolatey. Then to install software, at a command prompt, you just type the command for the package manager, the word install, and the package you want. For example, on a Mac, to get sqlite3, I might type:

```
brew install sqlite3
```

```
...or to upgrade...
```

```
brew upgrade sqlite3
```

For the coding part, I'd like you to write a simple command line application **RecordTester.java** to test properties of the record formats of Mysql and sqlite. You can assume the grader has Java 9 SE installed, but not necessarily any particular IDE or build suite. The grader will compile your code from the command line by switching into your folder, and entering the line:

```
javac RecordTester.java
```

The top of your program's class file should have any constants the grader needs to tweak to make a JDBC connection to a Mysql DB or sqlite Db. In the case of sqlite, you should as part of your homework folder have a empty file **test.sqlite** with an empty sqlite database file. Your program will be run from the command line with a command like:

```
java RecordTester dbms db table_name type num_rows num_columns
```

Here *dbms* is the type of database management system to connect to (either mysql (even if use MariaDB) or sqlite); *db* is the database to connect to; *table_name* is a table to create if it doesn't already exists, and drop and create if it already exists; *type* is a SQL attribute of INTEGER, BOOLEAN, or VARCHAR(n), num_rows is the number of rows of random data to be generated, num_cols is the number of columns of that type each row should have. An example set of inputs the grader might use is:

```
java RecordTester sqlite "test.sqlite" FOO "VARCHAR(7)" 100000 10
```

This should create in the sqlite database in the file test.sqlite, a table FOO which has 10 columns each of type CHAR(7), and then do 100000 inserts of random rows into this table.

Once you have written this program I want you to do experiments where you vary the SQL type and the number of rows and columns and for each type try to determine (a) how many bytes the DBMS needs to use for a record header and (b) how many bytes in a record is used to store the given type on average. To figure this out, you should locate the file on disk in which the table is stored and compare how the file size changes as you vary your programs parameters.

Carefully, write up each experiment you conduct and say what conclusion the experiment let's you draw. Put these write-ups in the file **Experiments.pdf**, which should also be included in your Hw1.zip file submitted.

Finally, to complete the description of the homework, your Hw1.zip should contain a readme.txt file in which the grader can find out anything else that might be needed to get you program to work as well as a list of all names the members of your homework group (at most three).

Point Breakdown

Written Problem 1 (1/2pt each sub-part)	2pts
Written Problem 2 (1/2pt each sub-part)	1pt
Written Problem 3 (1/2pt each sub-part)	1pt
RecordTester.java compiles and makes use of the command line arguments as described above	1pt
RecordTester succeeds in making a JDBC connection to the desired database	1pt
RecordTester creates the database table requested in the command line arguments and does the inserts as described by these arguments	1pt
Experiment.pdf with write-ups of experiments and results needed to determine header and attribute storage sizes for each of the types INTEGER, BOOLEAN, or VARCHAR(n) (1/2 pt each)	3pts
Total	10pts